HAROLD J. KUSHNER

ON DIFFUSIONS WITH CONTINUOUS DRIFT TERMS:
EXISTENCE, UNIQUENESS, STABILITY AND
APPROXIMATIONS

N71-22408 to
N71-22410

NASA CR-117746

CASE FILE
COPY

CENTER FOR DYNAMICAL SYSTEMS

ON DIFFUSIONS WITH DISCONTINUOUS DRIFT TERMS: EXISTENCE,
UNIQUENESS, STABILITY AND APPROXIMATIONS

by

Harold J. Kushner
Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island

This report consists of two papers dealing with properties

of diffusions with discontinuous drift terms.

1. Approximations to and Local Properties of
   Diffusions with Discontinuous Drift Terms.

2. Stability and Existence of Diffusions Dis-
   continuous or Rapidly Growing Drift Terms.

APPROXIMATIONS TO AND LOCAL PROPERTIES OF

DIFFUSIONS WITH DISCONTINUOUS CONTROLS

by

Harold J. Kushner[+]

Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island

APPROXIMATIONS TO AND LOCAL PROPERTIES OF
DIFFUSIONS WITH DISCONTINUOUS CONTROLS

Harold J. Kushner

1. Introduction.

The stochastic differential (Ito) equation (1)

(1) $dx = f(x,t,u(x,t))dt + \sigma(x,t)dz,$

where z_t is a Wiener process, is a common model of a variety of
stochastic control systems. Most past work has concentrated on the
case where σ and u satisfy a Lipschitz (or possibly a Holder)
condition in x, and f satisfies a similar condition in x and u.
This has been .so, owing to the existence of a very nice theory of
(1) under such conditions. Frequently, formal applications of
dynamic programming yield that the optimal control has surfaces of
discontinuity in x; for example, when f is linear in u, and the
cost does not depend on u. Recently, in an interesting paper,
Rishel [1] applied a transformation of Girsanov [2], to construct a
process of the form (1) where u is allowed to be merely bounded
and measurable, and proved some theorems concerning the relationship
between the formal dynamic programming equation for the cost, and
the optimal control.

Several questions remain open for the process constructed
in [1] or [2]. Since u is not necessarily uniformly Lipschitz, the

question of uniqueness remains. In Theorem 1, we show that, under reasonable conditions, the process constructed in [1] and [2] is a very natural solution to (1) (whether or not it is the unique solution), since it is the limit of the discrete time approximation, in the sense that the multivariate distributions converge. Uniqueness is then shown in Theorem 2. In Section 4, we replace the possibly discontinuous u by a sequence u_n which converges to u pointwise (except on the set of discontinuity of u), and prove that the costs of control converge to the cost for the discontinuous controls. Finally, in Section 5, we show that the local and growth properties of (1) are of the same type as for Lipschitz u. The control u is fixed, and the optimization problem is not treated here. Some approximation problems for the optimization problem will be treated in a companion paper.

2. Assumptions.

For a vector f and matrix σ define the norms $|f| = \sum_i |f_i|$, $|\sigma| = \sum_{i,j} |\sigma_{ij}|$, resp. Suppose that

(A1) the vectors $f^1(x,t), f^2(x,t)$, and matrix $\sigma(x,t)$ are Borel measurable in the pair x,t, satisfy a uniform Lipschitz condition (i.e., $|f(x,t) - f(y,t)| \leq K|x-y|$ for a real number K.) and a growth condition of the type $|f(x,t)|^2 \leq K(1+|x|^2)$. The matrix σ takes the form

$$(1) \qquad \sigma(x,t) = \begin{bmatrix} 0 & 0 \\ 0 & \hat{\sigma}(x,t) \end{bmatrix},$$

where $\hat{\sigma}(x,t)$ has a uniformly bounded inverse $\hat{\sigma}^{-1}(x,t)$. Note that this implies that $\hat{\sigma}^{-1}(x,t)$ satisfies a uniform Lipschitz condition in x.

Let z_t, $0 \leq t \leq T$, be a Wiener process with respect to the measure $P(\cdot)$, and define the Ito equation

$$(2) \qquad dx = \begin{pmatrix} dx^1 \\ dx^2 \end{pmatrix} = \begin{matrix} f^1(x,t)dt \\ f^2(x,t)dt + \hat{\sigma}(x,t)dz, \end{matrix} \quad 0 \leq t \leq T.$$

Assume that

(A2) The uncontrolled process x_t, given by (2), has a transition density $p(x,t; y,t+s)$ for $s > 0$. There are many important examples where this is the case. See Elliott [3], Kushner [4], Zakai [5] for some

frequently occuring types of examples.

Since the control $u(x,t)$ will appear in the form $\hat{f}(x,t) \equiv \hat{f}(x,t,u(x,t))$, only the properties of $\hat{f}$, as a function of x and t, will be important. For the most part we will deal only with the driving term $\hat{f}(x,t)$, and delete specific reference to the control.

(A3) Assume that the **vector** $\hat{f}(x,t)$ is bounded and Borel measurable in (x,t).

(A1) - (A3) are to be used throughout. It will occasionally be required that $\hat{f}(x,t)$ be continuous, except on a nice discontinuity set D_t, of Lebesgue measure zero. (Of course, we can write $\hat{f}(x,t)$ in the control-dependent form $\hat{f}(x,t,u(x,t))$, and translate (A4) into properties of $\hat{f}$ and u.) Thus, we write, for future use,

(A4) Let S_m denote a sphere whose center is the origin and with radius $M, N_\epsilon(A)$ an ϵ neighborhood of the set A, and $\mu(A)$, the Lebesgue measure of A. Suppose there is a set D_t (discontinuity set), so that for each M and t, $0 \leq t \leq T$, $\mu(N_\epsilon(D_t \cap S_M)) \to 0$ as$^+$ $\epsilon \to 0$. For each $\epsilon' > 0$, let there exist an $\epsilon > 0$ so that $|x-y| < \epsilon$ implies $|\hat{f}(x,t) - \hat{f}(y,t)| < \epsilon'$ (uniformly in x), provided that $x \notin N_\epsilon(D_t)$.

The control system to be considered has the form

$$(3) \qquad dx = \begin{pmatrix} dx^1 \\ dx^2 \end{pmatrix} = \begin{array}{l} f^1(x,t)dt \\ f^2(x,t)dt + \hat{f}(x,t)dt + \hat{\sigma}(x,t)d\tilde{z}, \end{array}$$

$$\equiv f(x,t)dt + \sigma(x,t)d\tilde{z},$$

$^+$Note that (A2) and (A4) imply that $P_{x_s}\{X_{t+s} \in N_\epsilon(D_{t+s})\} \to 0$ as $\epsilon \to 0$, where $P_{x_s}(\cdot)$ is the probability (for (2)) conditioned on x_s

where $\tilde{z}_t$ is a Brownian motion with respect to a measure $\tilde{P}(\cdot)$. Girsanov [2] and Rishel [1] construct the solution to (3) for a rather general class of $\hat{f}$, by a transformation of measures as follows. Let $\varphi(x,t)$ be a bounded vector-valued function of the same dimension as z_t.

In the rest of this paper x_t or $x_t(\omega)$ will always denote the random variables constructed in (2) or (6), namely the random variables which result when the usual Itô construction for (2) or (6) is used. However, the measures imposed on the x_t sequence may differ from usage to usage, and will be indicated where confusion may otherwise arise.[+] The initial condition $x_0 = x$ in (2) or (3) will be considered fixed, except where otherwise stated. Define (with respect to $P(\cdot)$)

$$(4) \qquad \zeta_0^T(\varphi) = \int_0^T \varphi'(x_s,s)dz_s - \frac{1}{2} \int_0^T \varphi'(x_s,s)\varphi(x_s,s)ds,$$

and the measure $\tilde{P}(\cdot)$ by $\tilde{P}(d\omega) = \exp \zeta_0^T(\varphi)\cdot P(d\omega)$. By [2], $\tilde{P}(\Omega) = 1$, and ([2], Theorem 1) $z_t - \int_0^t \varphi(x_s,s)ds \equiv \tilde{z}_t$ is a Wiener process with respect to $\tilde{P}(\cdot)$. Write (2) as

$$(5) \qquad dx_t = \begin{matrix} f^1(x_t,t)dt \\ f^2(x_t,t)dt + \hat{\sigma}(x_t,t)\varphi(x_t,t)dt + \hat{\sigma}(x_t,t)[dz-\varphi(x_t,t)dt] \end{matrix}$$

$$= \begin{matrix} f^1(x_t,t)dt \\ f^2(x_t,t)dt + \hat{\sigma}(x_t,t)\varphi(x_t,t)dt + \hat{\sigma}(x_t,t)d\tilde{z}_t. \end{matrix}$$

[+]Since all the measures to be imposed on the process x_t will be absolutely continuous with respect to $P(\cdot)$, the fact that x_t may only be unique w.p.1. will not be important.

Then $x_t(\omega)$ solves (5) with respect to $\tilde{P}(\cdot)$ and the Wiener process $\tilde{z}_t$. (i.e., the stochastic integral in (5) is computed using $\tilde{P}(\cdot)$.) By letting $\varphi = \hat{\sigma}^{-1}\hat{f}$ in (5) gives the desired equation (3). See [2], Theorem 1, for more details.

The following extension ([2], Theorem 1) will be useful Let the measurable (in (ω,t))$\sigma(\omega,t)$, $f^i(\omega,t)$, and $\varphi(\omega,t)$ be non-anticipative with respect to the z_t process, $0 < t < T$, $(P(\cdot))$, where $\varphi(\omega,t)$ is bounded and $E \int_0^T |\hat{\sigma}(\omega,t)|^2 < \infty$, $E \int_0^T |f^i(\omega,t)| dt < \infty$. Let x_t denote the unique random function which satisfies (relative to $P(\cdot)$)

$$(6) \qquad dx_t = \begin{array}{l} f^1(\omega,t)dt \\ f^2(\omega,t)dt + \hat{\sigma}(\omega,t)dz_t . \end{array}$$

Then $\tilde{z}_t \equiv z_t - \int_0^t \varphi(\omega,s)ds$ is a Wiener process with respect to the measure $\tilde{P}(d\omega) \equiv \exp \zeta_0^T(\varphi) \cdot P(d\omega)$ where $\zeta_0^T(\varphi)$ is defined by (4) with $\varphi(\omega,s)$ replacing $\varphi(x_s,s)$ and x_t satisfies (6). Then x_t also satisfies

$$(7) \qquad dx_t = \begin{array}{l} f^1(\omega,t)dt \\ f^2(\omega,t)dt + \hat{\sigma}(\omega,t)\varphi(\omega,t)dt + \hat{\sigma}(\omega,t)d\tilde{z}_t , \end{array}$$

relative to $\tilde{P}(\cdot)$.

Since the model (2) and its extension (3) are important in stochastic control theory, the purpose of this paper is to answer several questions which arise due to the non-constructive definition of the solution of (3).

(3) has a transition density if (2) does; indeed the

transition density is ([1], eqn (24))

$$
(8) \qquad \tilde{p}(x,t; \, y,t+s) = E\left[\exp \zeta_t^{t+s}(\varphi)\,\middle|\,x_{t+s}\right] p(x,t; \, y,t+s),
$$

where $p(x,t; \, y,t+s)$ is the transition density of (2).

3. Convergence of Discrete Time Approximations.

In this section, it will be shown that the process (3), as defined, is quite natural model for stochastic control purposes, being the limit (in the sense of convergence of the multivariate distributions) of the discrete time approximations to (3).

Lemma 1 ([2], Lemma 5). Let $\varphi_n(\omega,t)$ denote a sequence of uniformly bounded, measurable (in (ω,t) functions, non-anticipative with respect to the z_t process, for which $\zeta_0^T(\varphi_n) \to \zeta_0^T(\varphi)$ in probability $(P(\cdot))$, where $\varphi(\omega,t)$ is also bounded, measurable and non-anticipative. Then,

$$\int \left| \exp \zeta_0^T(\varphi_n) - \exp \zeta_0^T(\varphi) \right| P(d\omega) \to 0.$$

I.e., the measures $\tilde{P}^n(d\omega) = \exp \zeta_0^T(\varphi_n) \cdot P(d\omega)$ converge strongly to the measure $\tilde{P}(d\omega) = \exp \zeta_0^T(\varphi) \cdot P(d\omega)$. The expectations, $\tilde{E}^n g(\omega)$, of all bounded measurable functions $g(\omega)$ converge to $\tilde{E}g(\omega)$.

We add the following corollary which will be used frequently.

Corollary. Under the assumptions of Lemma 1, for any bounded measurable function $g(\omega)$, we have[+] $\tilde{E}^n g(\omega) \to \tilde{E}g(\omega)$. Let $\varphi_n(\omega,t)$ denote a sequence of uniformly bounded, measurable functions which are non-anticipative with respect to the z_t process $(P(\cdot))$. Let $\varphi_n(\omega,t) \to \varphi(\omega,t)$ in probability $(P(\cdot))$ for almost all t in $[0,T]$.

[+] $\tilde{E}^n$ denotes expectation with respect to $\tilde{P}^n(d\omega) = P(d\omega)\exp \zeta_0^T(\varphi_n)$.

Then $\zeta_0^T(\varphi_n) \to \zeta_0^T(\varphi)$ in probability.

Proof. The first assertion follows directly from the Lemma. For the proof of the second assertion, note that $E|\varphi_n(\omega,t) - \varphi(\omega,t)|^4$ are uniformly bounded, and converge to zero as $n \to \infty$ for almost all t in $[0,T]$. This implies that

$$E[\int_0^T (\varphi_n(\omega,t) - \varphi(\omega,t))'dz_t]^2 \to 0$$

$$E[\int_0^T [\varphi_n'(\omega,t)\varphi_n(\omega,t) - \varphi'(\omega,t)\varphi(\omega,t)]^2 \to 0. \qquad Q.E.D.$$

The finite difference approximation. For some Wiener process $\tilde{z}_t$, define the discrete parameter process, with $x_0^\Delta = x_0$,

$$(9) \quad x_{n+1}^\Delta = x_n^\Delta + \begin{pmatrix} \int_{n\Delta}^{n\Delta+\Delta} f^1(x_n^\Delta,s)ds \\ \int_{n\Delta}^{n\Delta+\Delta} f^2(x_n^\Delta,s)ds \end{pmatrix} + \int_{n\Delta}^{n\Delta+\Delta} \hat{f}(x_n^\Delta,s)ds + \int_{n\Delta}^{n\Delta+\Delta} \hat{\sigma}(x_n^\Delta,s)d\tilde{z}_s$$

and the interpolation of (9), where $x_t^\Delta = x_n^\Delta$ for $n \leq t < n\Delta + \Delta$,

$$(10) \quad dx_t^\Delta = \begin{pmatrix} f^1(x_t^\Delta,t)dt \\ f^2(x_t^\Delta,t)dt \end{pmatrix} + \hat{f}(x_t^\Delta,t)dt + \hat{\sigma}(x_t^\Delta,t)d\tilde{z}_t.$$

The multivariate distributions of the discrete parameter

process (9) and of its interpolation (10), do not depend on the specific Wiener process used, nor on the method of construction. Hence we may construct the processes (9) and (10) in any convenient manner. Now define the 'uncontrolled' discrete parameter process

$$(11) \qquad x_{n+1}^{\Delta} = x_n^{\Delta} + \begin{pmatrix} \int_{n\Delta}^{n\Delta+\Delta} f^1(x_n^{\Delta}, s)ds \\[2em] \int_{n\Delta}^{n\Delta+\Delta} f^2(x_n^{\Delta}, s)ds \end{pmatrix} + \int_{n\Delta}^{n\Delta+\Delta} \hat{\sigma}(x_n^{\Delta}, s)dz_s,$$

where z_t is a Wiener process $(P(\cdot))$, and its interpolation (12), where $x_t^{\Delta} = x_n^{\Delta}$ for $n\Delta \leq t < n\Delta + \Delta$,

$$(12) \qquad dx_t^{\Delta} = \begin{pmatrix} f^1(x_t^{\Delta}, t)dt \\[1.5em] f^2(x_t^{\Delta}, t)dt \end{pmatrix} + \hat{\sigma}(x_t^{\Delta}, t)dz_s$$

Since the z_t and x_t^{Δ} in (11) - (12) are a Wiener and an Itô process, resp., with respect to the measure $P(\cdot)$, the process defined by

$$\tilde{z}_t^{\Delta} \equiv z_t - \int_0^t \hat{\sigma}^{-1}(x_s^{\Delta}, s)\hat{f}(x_s^{\Delta}, s)ds$$

is a Wiener process with respect to the measure

$$\widetilde{P}^{\triangle}(d\omega) = P(d\omega)\exp \zeta_0^T(\varphi_{\triangle})$$

$$\zeta_0^T(\varphi_{\triangle}) = \int_0^T [\hat{\sigma}^{-1}(x_s^{\triangle},s)\hat{f}(x_s^{\triangle},s)]'\,dz_s$$

$$-\frac{1}{2}\int_0^T [\hat{\sigma}^{-1}(x_s^{\triangle},s)\hat{f}(x_s^{\triangle},s)]'\,[\hat{\sigma}^{-1}(x_s^{\triangle},s)\hat{f}(x_s^{\triangle},s)]\,ds$$

where $x_t^{\triangle}$ is the process constructed in (11) (with respect to $P(\cdot)$). Now, the $x_t^{\triangle}$ process of (11) also satisfies (10) with respect to the Wiener process $z_t^{\triangle}$. The multivariate distributions of the processes (9) - (10) constructed in this way must agree with the distributions corresponding to any other construction of the processes (9) - (10).

Theorem 1. Assume (A1) - (A4). Then the multivariate distributions of (10) converge to those of (3).

Proof. Let the symbol E denote expectation with respect to $P(\cdot)$. In view of the foregoing discussion, we need only prove the convergence assertion (for each set of real vectors $\lambda_1,\ldots,\lambda_m$)

$$\int \exp(i\sum_{j=1}^m \lambda_j' x_{t_j}^{\triangle})\widetilde{P}^{\triangle}(d\omega) \equiv \widetilde{F}^{\triangle}(\lambda_1,\ldots,\lambda_m)$$

$$= E\,\exp(i\sum_{j=1}^m \lambda_j' x_{t_j}^{\triangle} + \zeta_0^T(\varphi_{\triangle}))$$

$$\to E\,\exp(i\sum_{j=1}^m \lambda_j' x_{t_j} + \zeta_0^T(\varphi)) = \int \exp(i\sum_{j=1}^m \lambda_j' x_{t_j})\widetilde{P}(d\omega)$$

$$\equiv \widetilde{F}(\lambda_1,\ldots,\lambda_m),$$

where x_t is the Itô process (2), and φ is given by (4). Recall that $E|x_t^\Delta - x_t|^4 \to 0$.

Next, we prove that the measures $\tilde{P}^\Delta(\cdot)$ converge strongly to $\tilde{P}(\cdot)$.

In view of Lemma 1 and its corollary, we only need to show that

$$\varphi_\Delta(\omega, t) \equiv \hat{\sigma}^{-1}(x_t^\Delta, t)\hat{f}(x_t^\Delta, t) \xrightarrow{P} \hat{\sigma}^{-1}(x_t, t)\hat{f}(x_t, t) \equiv \varphi(\omega, t).$$

Since $\sigma^{-1}(x, t)$ satisfies a uniform Lipschitz condition in x, and $E|x_t^\Delta - x_t|^4 \to 0$, we need only show that, for almost all t in $[0, t]$,

$$\hat{f}(x_t^\Delta, t) \xrightarrow{P} \hat{f}(x_t, t).$$

Equivalently, we show that for each fixed $\delta > 0$, $\epsilon' > 0$,

$$(13) \qquad P\{|\hat{f}(x_t^\Delta, t) - \hat{f}(x_t, t)| > \epsilon'\} < \delta.$$

for sufficiently small Δ. Let ϵ and ϵ' correspond by (A4), where $\epsilon' \to 0$ as $\epsilon \to 0$, and let $\epsilon_1 \le \epsilon$. The left side of equation (13) can be bounded above by the sum

$$P\{x_t \in N_{\epsilon_1}(D_t)\} + P\{|x_t^\Delta - x_t| > \epsilon_1] \equiv \delta_1 + \delta_s,$$

Let ϵ_0 be small enough so that $\delta_1 < \delta/2$ for $\epsilon_1 \le \epsilon_0$. Let ϵ_1 denote the minimum of ϵ_0 and the ϵ which corresponds to ϵ' by (A4).

Then $\delta_2 < \delta/2$ for sufficiently small Δ.

Thus (13) holds for arbitrary ϵ' and δ, and sufficiently small Δ.

By ([1], Lemma 1), we have

$$(14) \qquad E \exp \alpha \zeta_0^T(\varphi) \leq \exp \frac{1}{2}(\hat{\alpha}^2 - \alpha)Tc^2 = B_\alpha$$

where c is proportional to the bound on $|\varphi(\omega,t)|$. Let $g_\Delta(\omega) \xrightarrow{P} g(\omega)$. Then

$$(15) \quad \tilde{E}^\Delta |g_\Delta - g| = E \exp \zeta_0^T(\varphi_\Delta)|g_\Delta - g| \leq E^{1/2} \exp 2\zeta_0^T(\varphi_\Delta) E^{1/2} |g_n - g|^2 \to 0,$$

since the φ_Δ have a common bound. Since $\exp i \sum_{j=1}^m \lambda_j' x_{t_j}^\Delta \xrightarrow{P}$

$\exp i \sum_{j=1}^m \lambda_j' x_{t_j}$, letting $g_\Delta = \exp i \sum_{j=1}^m \lambda_j' x_{t_j}^\Delta$, and $g = \exp i \sum_{j=1}^m \lambda_j' x_{t_j}$,

and using the strong convergence of $\tilde{P}^\Delta(\cdot)$ to $\tilde{P}(\cdot)$, gives

$$\tilde{F}^\Delta(\lambda_1, \ldots, \lambda_n) = \tilde{E}^\Delta g_\Delta = \tilde{E}^\Delta g + \tilde{E}^\Delta(g_\Delta - g) \to \lim \tilde{E}^\Delta g = \tilde{E}g = \tilde{F}(\lambda_1, \ldots, \lambda_n). \quad \text{Q.E.D.}$$

Theorem 2. Assume (A1) - (A4). Then the solution x_t of (3) is unique, in the sense that if x_t^i, $i = 1, 2$, solve (3) with the Wiener processes $\tilde{z}_t^i$, the multivariate distributions of the process x_t^i, $i = 1, 2$, are the same.

Proof. Suppose that the Itô and Wiener processes x_t and $\tilde{z}_t$ solve (3) with respect to the measure $\tilde{P}(\cdot)$. Define the measure

$$P(d\omega) = \tilde{P}(d\omega)\exp \zeta_0^T (-\tilde{\varphi})$$

$$-\tilde{\varphi}_s = -\hat{\sigma}^{-1}(x_s,s)\hat{f}(x_s,s)$$

$$\zeta_0^T(-\tilde{\varphi}) = -\int_0^T \tilde{\varphi}_s' d\tilde{z}_s - \frac{1}{2} \int_0^T \tilde{\varphi}_s'\tilde{\varphi}_s ds.$$

The stochastic integral in $\zeta_0^T(-\tilde{\varphi})$ is defined with respect to the measure $\tilde{P}(\cdot)$. But, since $P(\cdot)$ and $\tilde{P}(\cdot)$ are absolutely continuous with respect to each other, the random variable $\int_0^T \tilde{\varphi}'(\omega,s)d\tilde{z}_s$ is defined uniquely w.p.1. with respect to $P(\cdot)$ also. Now $z_t \equiv \tilde{z}_t + \int_0^t \hat{\sigma}^{-1}(x_s,s)\hat{f}(x_s,s)ds$ is a Wiener process $(P(\cdot))$ and (2) is an Itô process (with respect to the just constructed $z_t,P(\cdot)$). Starting with the constructed triple $P(\cdot)$, z_t,x_t, we can re-obtain the original $\tilde{P}(\cdot),\tilde{z}_t,x_t$ by the usual Girsanov transformation on (2). I.e., there is some Wiener process z_t (with respect to $P(\cdot)$), and Itô process (2), so that $\tilde{z}_t$ and the solution to (3) can be constructed by the Girsanov transformation on z_t and the x_t given by (2). **Since** the multivariate distributions of the discrete approximations (10) do not depend on the Wiener process, and converge to those for (3) for any solution to (3) which is obtainable by a Girsanov transformation, it is clear that the multivariate distributions of (3) are unique. Q.E.D.

4. Approximations of discontinuous controls, and the corresponding costs.

Define the processes (16) and (17) constructed by the Girsanov transformation on (2) with measures

$$P(d\omega)\exp \zeta_0^T(\varphi_n), \quad P(d\omega)\exp \zeta_0^T(\varphi),$$

resp., where $\varphi_n(x_t, t) = \hat{\sigma}^{-1}(x_t, t)\hat{f}_n(x_t, t)$

$$(16) \qquad dx_t = \begin{pmatrix} f^1(x_t, t)dt \\ \\ f^2(x_t, t)dt \end{pmatrix} + \hat{f}_n(x_t, t)dt + \hat{\sigma}(x_t, t)d\tilde{z}_t^n$$

$$(17) \qquad dx_t = \begin{pmatrix} f^1(x_t, t)dt \\ \\ f^2(x_t, t)dt \end{pmatrix} + \hat{f}(x_t, t)dt + \hat{\sigma}(x_t, t)d\tilde{z}_t.$$

Let $\tilde{E}^n$ and $\tilde{E}$ denote the expectation corresponding to (16) and (17), resp.

$\hat{f}_n(x, t)$ is to be an approximation to $\hat{f}(x, t)$ (or, equivalently, $u_n(x, t)$ is an approximation to $u(x, t)$, and $\hat{f}_n(x, t) \equiv \hat{f}(x, t, u_n(x, t)))$. The next two theorems consider the convergence of costs as $u_n \to u$ or $\hat{f}_n \to \hat{f}$. The results are of interest, since we may desire to approximate the discontinuous control $u(x, t)$ by a smooth control $u_n(x, t)$ which is smooth in x and where $u_n(x, t) \approx u(x, t)$ except in a neighborhood of the discontinuity set D_t.

Theorem 3. Assume (A1) - (A3). Let $k(x,t)$ and $b(x)$ be bounded and Borel measurable. Let $\hat{f}_n(x,t) \to \hat{f}(x,t)$ for each $x \notin D_t$, where $\hat{f}_n$ and $\hat{f}$ are uniformly bounded, and Borel measurable. Then

$$(18) \qquad \tilde{E}^n \int_0^T k(x_s,s)ds + \tilde{E}^n b(x_T) \to \tilde{E} \int_0^T k(x_s,s)ds + \tilde{E}b(x_T).$$

Let $u_n(x,t)$ and $u(x,t)$ be Borel measurable. Let $k(x,s,u)$ be bounded, Borel measurable, and let $k(x,s,u_n(x,s)) \to k(x,s,u(x,s))$, as $n \to \infty$, except for $x \in D_t$, for $t > 0$. Then (18) holds true if $k(x,s,u)$ replaces $k(x,s)$.

Proof. Let $g(x)$ be bounded and Borel measurable. To prove the first assertion, we only need to show that

$$(19) \qquad \tilde{E}^n g(x_t) \to \tilde{E}g(x_t), \quad t > 0.$$

If $\zeta_0^T(\varphi_n) \xrightarrow{P} \zeta_0^T(\varphi)$, then, by Lemma 1, the measures $\tilde{P}^n(\cdot)$ converge strongly to $\tilde{P}(\cdot)$, and (19) holds.

But, as in the proof of Theorem 1, $\zeta_0^T(\varphi_n) \xrightarrow{P} \zeta_0^T(\varphi)$ if

$$\hat{f}_n(x_t,t) \xrightarrow{P} \hat{f}(x_t,t), \quad t > 0,$$

which holds since $\hat{f}_n(x,t) \to \hat{f}(x,t)$ except for $x \in D_t$, where $P\{x_t \in D_t\} = 0$.

Continuing to the second assertion, we use the argument

which led from (15) to the end of the proof of Theorem 1. We must

show that

$$\tilde{E}^n k(x_t,t,u_n(x_t,t)) \to \tilde{E}k(x_t,t,u(x_t,t)).$$

Since, by the first part of the theorem,

$$\tilde{E}^n k(x_t,t,u(x_t,t)) \to \tilde{E}k(x_t,t,u(x_t,t)),$$

we need only show that

(20) $$\tilde{E}^n |k(x_t,t,u_n(x_t,t)) - k(x_t,t,u(x_t,t))| \to 0.$$

But, hypothesis, the integrand in the left side of (20) converges to zero

P(·)). Then an argument identical to (15) gives the conclusion (20). Q.E.D.

The next theorem concerns the cost approximation for a process

which is stopped, not at fixed time T, but at the random time

$\tau = \inf \{t: x_t \notin R\}$, where R is a fixed bounded open set. Note that

$\min [\tau(\omega),T]$ depends on the path $x_t(\omega)$, where x_t is given by (2). The

value of the random variable $\min [\tau(\omega),T]$ is the same for the transformed

process (3) (with $\hat{f}_n$ or $\hat{f}$ used) as for (2), since the process paths x_t

are the same. Only the measures are changed. Note that, since the $\hat{f}_n$ are

uniformly bounded, there is a constant $K < \infty$ for which

$$\tilde{E}^n \tau \leq K < \infty, \quad \tilde{E}\tau \leq K.$$

This is implied by the fact that $\sigma(x)\sigma'(x) \geq \alpha I$, where I is the identity

matrix, for some $\alpha > 0$.

Theorem 4. <u>Assume the conditions of the second paragraph of Theorem 3. Then</u>

(21)
$$\tilde{E}^n \int_0^\tau k(x_s, s, u_n(x_s, s)) ds + \tilde{E}^n b(x_\tau) \to$$
$$\tilde{E} \int_0^\tau k(x_s, s, u(x_s, s)) ds + \tilde{E} b(x_\tau)$$

<u>as</u> $n \to \infty$.

Proof. Fix $T < \infty$. First we show that (21) holds for $T \cap \tau \equiv \min(T, \tau)$ replacing τ. Since the measures (for the processes on $[0, T]$ for each $T < \infty$) $\tilde{P}^n(\cdot)$ converge to $\tilde{P}(\cdot)$ strongly, the expectations of any bounded and measurable function of ω converge, as $n \to \infty$. Let $I_\tau(s) = 1$ if $\tau \geq s$ and 0 if $\tau < s$. Thus

(22)
$$\tilde{E}^n b(x_{\tau \cap T}) = \tilde{E}^n b(x_{\tau(\omega) \cap T(\omega)}) \to \tilde{E} b(x_{\tau \cap T})$$
$$\tilde{E}^n k(x_s, s, u(x_s, s)) I_\tau(s) \to \tilde{E} k(x_s, s, u(x_s, s)) I_\tau(s),$$

for $s < T$ and, as in the proof of Theorems 1 or 3

(23)
$$\tilde{E}^n |k(x_s, s, u(x_s, s)) - k(x_s, s, u_n(x_s, s))| I_\tau(s) \to 0.$$

Thus (see proof of Theorem 3)(21) holds with $T \cap \tau$ replacing τ.

Next, since $\tilde{E}^n \tau < K < \infty$, $\tilde{E} \tau < K < \infty$, we have

$$\widetilde{E}^n \int_{T \cap \tau}^{\tau} |k(x_s, s, u_n(x_s, s))| \, ds \to 0$$

(24)

$$\widetilde{E}^n |b(x_\tau) - b(x_{\tau \cap T})| \to 0$$

as $T \to \infty$, uniformly in n. Equation (21) follows from (22) - (24). Q.E.D.

5. Local properties of (3).

In this Section we give some of the local properties of (3), as defined by the Girsanov transformation. Inde d, (3) has virtually all the properties which it would have if $\hat{f}$ were Lipschitz in x.

Theorem 5. Assume (A1) - (A4). Let x_t^n and x_t denote the processes given by (3), with initial conditions $x_0^n = x^n$ and $x_0 = x$. (We use $\tilde{E}^n$ and $\tilde{E}$ for the corresponding expectation operators.) Then the multivariate distributions of x_t^n converge to those of x_t as $x^n \to x$.

Proof. In the proof, x_t^n and x_t denote the random variables defined by (2), corresponding to the initial conditions x^n and x, resp. We must show that, for any set of real vectors $\lambda_1, \ldots, \lambda_m$,

(25)
$$\tilde{E}^n \exp i \sum_{j=1}^m \lambda_j' x_{t_j}^n = E \exp [\zeta_0^T(\varphi_n) + i \sum_{j=1}^m \lambda_j' x_{t_j}^n] \to$$
$$E \exp [\zeta_0^T(\varphi) + i \sum_{j=1}^m \lambda_j' x_{t_j}] = \tilde{E} \exp i \sum_{j=1}^m \lambda_j' x_{t_j}$$

where $\zeta_0^T(\varphi_n)$ is (4) with x_s^n replacing x_t. Now

$$\exp i \sum_{j=1}^m \lambda_j' x_{t_j}^n \xrightarrow{P} \exp i \sum_{j=1}^m \lambda_j' x_{t_j} .$$

As in the proof of Theorem 3, we only need to show that

$\zeta_0^T(\varphi^n) \xrightarrow{P} \zeta_0^T(\varphi)$, for this implies that $\tilde{P}^n(\cdot) \to \tilde{P}(\cdot)$ strongly (see proof of Theorem 3). But, $\zeta_0^T(\varphi^n) \xrightarrow{P} \zeta_0^T(\varphi)$ if $\hat{f}(x_t^n, t) \xrightarrow{P} \hat{f}(x_t, t)$, as $x^n \to x$. Now, $(\epsilon, \epsilon'$ are defined by (A4) and $\epsilon_1 \leqq \epsilon)$

$$P\{|\hat{f}(x_t^n, t) - \hat{f}(x_t, t)| > \epsilon'\} \leq P\{|x_t^n - x_t| > \epsilon_1\} +$$

$$P\{x_t \in N_{\epsilon_1}(D_t)\} \equiv \delta_1 + \delta_2.$$

Choose $\epsilon' > 0$ and $\delta > 0$. Choose ϵ_0 so that $\delta_2 < \delta/2$, for $\epsilon_1 < \epsilon_0$. (This is possible since x_t has a density.) Let ϵ_1 denote the minimum of ϵ_0 and the ϵ corresponding to ϵ' by (A4). Then for large enough n, $\delta_1 < \delta/2$ which proves the assertion. Q.E.D.

Theorem 6. K _and_ K_i _denote real numbers._ _Fix_ T, _and let_ $x_0 = x$. _Assume_ (A1) - (A4). _Then, for_ $0 < t < T$,

$$(26) \qquad \tilde{E}|x_t|^2 \leq K_1(|x|^2 + 1),$$

$$(27) \qquad \tilde{E} \sup_{0 < t < T} |x_t|^2 \leq K_2(|x|^2 + 1)$$

$$(28) \qquad \tilde{E}|x_t - x|^2 \leq K_3 t(|x|^2 + 1).$$

For $x \notin D_0$,

$$(29) \quad \left| \tilde{E}x_t - x - \left\{ \begin{array}{c} \int_0^t f^1(x,s)ds \\ \\ \int_0^t f^2(x,s)ds \end{array} \right\} - \int_0^t \hat{f}(x,s)ds \right| \le o(t)(1+|x|^2)^{1/2}$$

where $o(t)/t \to 0$ as $t \to 0$ uniformly in x, for $0 < t \le T$.

$$(30) \quad \left| E(x_t-x)(x_t-x)' - \int_0^t \sigma(x,s)\sigma'(x,s)ds \right| \le K_4(|x|^2+1)t^2.$$

Let $\mathscr{B}_r$ denote the minimum σ-algebra which measures $x_s, s \le r$. Then, w.p.l.,

$$(31) \quad \left| \tilde{E}(x_{t+r} | \mathscr{B}_r) - x_r - \left\{ \begin{array}{c} \int_r^{t+r} f^1(x_r,s)ds \\ \\ \int_r^{t+r} f^2(x_r,s)ds \end{array} \right\} - \int_r^{t+r} \hat{f}(x_r,s)ds \right|$$

$$\le o(t)(1+|x_r|^2)^{1/2}$$

where $o(t)$ satisfies the property above (30). Also, w.p.l.,

$$\left| \tilde{E}\{(x_{t+r}-x_r)(x_{t+r}-x_r)' | \mathscr{B}_r\} - \int_r^{t+r} \sigma(x_r,s)\sigma'(x_r,s)ds \right|$$

$$(32)$$

$$\le K_5(|x_r|^2+1)t^2.$$

Proof. Proof of (26). Note that all moments of (3) are finite since all moments of (2) are finite and

$$\tilde{E}^2 |x_t|^m \leq E|x_t|^{2m} E \exp 2\zeta_0^T(\varphi) < \infty.$$

Using the estimates $\tilde{E} \ |\int_0^t \sigma(x_s,s)d\tilde{z}_s|^2 \leq \tilde{E} \int_0^t |\sigma(x_s,s)|^2 ds$ and $(\int_0^t |f| ds)^2 \leq t \int_0^t f^2 ds$, the boundedness of $|\hat{f}|$ and the growth condition (Al) in σ and f^i on equation (3) yields

$$(33) \quad \tilde{E}|x_t|^2 \leq K_6|x|^2 + K_6 \int_0^t (1+\tilde{E}|x_s|^2)ds + K_6(t+t^2), \quad 0 \leq t \leq T.$$

(33), together with the finiteness of $\tilde{E}|x_t|^2$ implies (26). The proof of (27) is rather standard and is similar to that of (26), the main difference being the use of the martingale estimate

$$\tilde{E} \max_{0 < t < T} |\int_0^t \sigma(x_s,s)d\tilde{z}_s|^2 \leq 2 \int_0^t \tilde{E}|\sigma(x_s,s)|^2 ds,$$

and the details are omitted.

The proof of (28) is essentially the same as that for the case where $\hat{f}$ is Lipschitz in x, and is omitted.

Proof of (29). We have

$$(34a) \qquad \widetilde{E} \int_0^t |f^i(x_s,s) - f^i(x,s)| \, ds \le K \int_0^t \widetilde{E} |x_s - x| \, ds \le K_7 \cdot t^{3/2}(1+|x|^2)^{1/2}.$$

Also

$$I_s \equiv \widetilde{E}|\hat{f}(x_s,s) - \hat{f}(x,s)| = \int_{|x_s-x| \ge \epsilon} |\hat{f}(x_s,s) - \hat{f}(x,s)| \widetilde{P}(d\omega)$$

$$+ \int_{|x_s-x| < \epsilon} |\hat{f}(x_s,s) - \hat{f}(x,s)| \widetilde{P}(d\omega)$$

$$\equiv I_{1s} + I_{2s}.$$

Note that, by the first part of (A4), $x \not\in N_\epsilon(D_0)$ for sufficiently small ϵ, say $\epsilon \le \epsilon_0$ (since $x \not\in D_0$). Also, using the boundedness of $\hat{f}$ and Chebyshev's inequality, and (28),

$$I_{1s} \le K_8 P\{|x_s-x| \ge \epsilon\} \le K_9 \frac{s^{1/2}(1+|x|^2)^{1/2}}{\epsilon}$$

$$I_{2s} \le \epsilon',$$

where (see (A4)) we define $\epsilon' = \epsilon'(\epsilon) = \sup_{\substack{|x-y|<\epsilon \\ x \not\in N_\epsilon(D_s)}} |\hat{f}(x,s) - \hat{f}(y,s)|$ and

$\epsilon' \to 0$ as $\epsilon \to 0$, uniformly in x, for $x \not\in N_{\epsilon_0}(D)$, for any fixed $\epsilon_0 > 0$.

Let $\epsilon = s^{1/4}$. Then

$$(34b) \qquad \int_0^t I_s \le K_{10} t^{5/4}(1+|x|^2)^{1/2} + o(t),$$

where $o(t)/t \to 0$ as $t \to 0$ uniformly in x, for $x \not\in N_{\epsilon_0}(d)$, for any $\epsilon_0 > 0$. (34a) and (34b) imply (29), for sufficiently small t.

The proofs of (30) and (32) are the same for the case where $\hat{f}$ is Lipschitz - if $|\hat{f}|$ is replaced by its supremum in all estimates. (The drift terms do not effect the result as long as the discontinuous term is bounded.)

Proof of (31). More generally than (28), it can be shown that, w.p.l.,

$$\tilde{E}\{|x_{t+r} - x_r|^2 | \mathscr{B}_r\} \leq K_{11} \cdot t(|x_r|^2 + 1), \quad 0 \leq t \leq T.$$

It follows from this and Schwarz's inequality that

$$(35) \quad \tilde{E}\{\int_r^{t+r} |f^i(x_s,s)-f^i(x_r,s)|\,ds | \mathscr{B}_r\} \leq t^{1/2}\,\tilde{E}[\int_r^{t+r} |f^i(x_s,s)-f^i(x_r,s)|^2\,ds | \mathscr{B}_r]^{1/2} \leq$$

$$\leq K_{12}t^{3/2}(1+|x_r|^2)^{1/2}.$$

Also, for $s > r$,

$$I_s \equiv \tilde{E}\{|\hat{f}(x_s,s) - \hat{f}(x_r,s)| | \mathscr{B}_r\} = \int |\hat{f}(x_s,s) - \hat{f}(x_r,s)|\tilde{p}(x_r,r; x_s,s)dx_s$$

where the integral is evaluated at $x_r = x_r(\omega)$, and $\tilde{p}(x,r; y,s)$ is the transition density. We may bound I_s by

$$\int_{\substack{|x_s-x_r|<\epsilon_s \\ x_r \notin N_{\epsilon_s}(D_r)}} |\hat{f}(x_s,s) - \hat{f}(x_r,s)|\tilde{p}(x_r,r; x_s,s)dx_s$$

$$+ \int_{|x_s-x_r|\geq\epsilon_s} |\hat{f}(x_s,s) - \hat{f}(x_r,s)|\tilde{p}(x_r,r; x_s,s)dx_s$$

$$+ \int_{x_r \in N_{\epsilon_s}(D_r)} |\hat{f}(x_s,s) - \hat{f}(x_r,s)|\tilde{p}(x_r,r; x_s,s)dx_s \equiv I_{1s} + I_{2s} + I_{3s}$$

where

$$I_{1s} \leq K_{13} \epsilon'_s$$

$$I_{2s} \leq \frac{K_{13} s^{1/2} (1+|x_r|^2)^{1/2}}{\epsilon_s}$$

$$I_{3s} \leq K_{13} \epsilon''(\epsilon_s) \equiv K_{13} \tilde{P}\{|x_r| \in N_{\epsilon_s}(D_r)\} \to 0, \text{ as } \epsilon_s \to 0,$$

where (see (A4)) $|x-y| < \epsilon_s$ and $x \notin N_{\epsilon_s}(D_t)$ imply that

$|\hat{f}(x,t) - \hat{f}(y,t)| < \epsilon'_s$ where $\epsilon'_s \to 0$ as $\epsilon_s \to 0$. Letting $\epsilon_s = s^{1/4}$

gives $I_{1s} + I_{2s} + I_{3s} = \epsilon'''(s)[1+|x_r|^2]^{1/2}$, where $\epsilon'''(s) \to 0$

as $s \to 0$, uniformly in x_r. Now

$$\int_0^t \epsilon'''(s)ds = o(t)$$

uniformly in x_r, which establishes (31). Q.E.D.

References

[1]   R. W. Rishel, "Weak Solutions of the Partial Differential Equation of Dynamic Programming", submitted to SIAM Journal on Control.

[2]   I. V. Girsanov, "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures", Theory of Probability and its Applications, 5, 1960, pp. 285-301, translation.

[3]   D. L. Elliot, "Behavior of Controllable Nonlinear Systems Driven by White Noise", Second Hawaiian Conference on System Science, January 1969.

[4]   H. J. Kushner, "The Cauchy Problem for a Class of Degenerate Parabolic Equations, and Asymptotic Properties of Related Diffusion Processes", J. Differential Equations, 6, 1969, pp. 209-231.

[5]   M. Zakai, "A Lyapunov Criteria for the Existence of Stationary Probability Distributions for Systems Perturbed by Noise", SIAM Journal on Control, 7, 1969, pp. 390-397.

STABILITY AND EXISTENCE OF DIFFUSIONS WITH
DISCONTINUOUS OR RAPIDLY GROWING DRIFT TERMS[+]

by

H. J. Kushner

Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island

Abstract

Stochastic differential equations whose drift terms do not satisfy the usual (Itô) Lipschitz or linear growth conditions in the state occur frequently as models in stochastic control theory. Local stability properties are useful for proving global existence for ordinary differential equations whose right hand sides grow too fast or are not Lipschitz in the state. Here, we use a local stochastic stability property to prove global existence, stability, ergodicity, the strong Markov and other properties, for a class of diffusions which occur frequently as models.

1. Introduction

For a vector $x = \{x_i\}$ and matrix $\sigma = \{\sigma_{ij}\}$, define the Euclidean norms $|x|^2 = \sum_i x_i^2$, $|\sigma|^2 = \sum_{i,j} \sigma_{ij}^2$, resp. Consider the homogeneous[+] Itô stochastic differential equation

$$(1) \qquad dx = f(x)dt + \sigma(x)dz, \qquad t \geq 0$$

where $\sigma(\cdot)$ satisfies growth and Lipschitz conditions of the types[++]

$$(2a) \qquad |\sigma(x)|^2 \leq K(1+|x|^2)$$

$$(2b) \qquad |\sigma(x) - \sigma(y)| \leq K(1+|x|),$$

and $z(t)$ is a normalized vector valued Wiener process. If

$$(3a) \qquad |f(x)|^2 \leq K(1+|x|^2)$$

$$(3b) \qquad |f(x) - f(y)| \leq K|x-y|$$

then the Ito existence theory is applicable to (1) and the stability properties can be discussed [1]. If (3b) holds locally, but (3a) is violated, a 'local' stability property([1], Theorem 8, Chapter 2) ensures the existence of a solution to (1) for all $t > 0$.

[+] The homogeneity condition is not essential, except in Section 4.

[++] K and K_i always denote real numbers; their value may change from usuage to usuage.

2-2

Recent investigations [2-5] have studies an important class of equations (1), where $f(\cdot)$ is allowed some discontinuities. Rewrite (1) in the form (x^1 and x^2 are vectors).

$$(4) \qquad dx = \begin{pmatrix} dx^1 \\ dx^2 \end{pmatrix} = \begin{matrix} f^1(x)dt \\ f^2(x)dt + \hat{f}(x)dt + \hat{\sigma}(x)dz \end{matrix}$$

where we assume that the f^i and $\hat{\sigma}$ satisfy (3) and (2), respectively, and $\hat{\sigma}(x)$ has a uniformly bounded inverse. (Thus $\hat{\sigma}^{-1}(x)$ satisfies (2).), but $\hat{f}(\cdot)$ does not necessarily satisfy (3). In the sequel, we prove existence, uniqueness, and other properties of (4), when neither (3a) nor (3b) necessarily holds, but a 'local' stability property obtains, and also treat the problems of asymptotic stability, the existence of a unique invariant measure and the convergence of the measures of (1) to the invariant measure.

Diffusions of the type (4) occur frequently in control applications. Consider, for example, a 'white noise' driven n'th order differential equation where $\hat{f}$ is a 'bang-bang' control taking the values $\{+1,-1\}$, or which may be discontinuous on a smooth 'switching curve', and tend to infinity in certain directions. Also models such as

$$dx = \begin{matrix} dx_1 = x_2dt \\ dx_2 = -(x_1+x_1^3)dt + \sigma dz \end{matrix}$$

are sometimes used, and the existence, and asymptotic character of the corresponding measures are of interest.

2. Mathematical Preliminaries

Assume

(C1) f^i and $\hat{\sigma}$ satisfy (3) and (2), respectively, and $\hat{\sigma}^{-1}(x)$ is uniformly bounded. $\hat{f}(\cdot)$ is a vector valued Borel function of x which is bounded in any compact set.

(C2) The process (5) has a transition density $p(x; t,y)$.

(C3) (A condition on the discontinuities of $\hat{f}$.) Let S_m denote a shpere of radius m, whose center is the origin. Let $N_\epsilon(A)$ denote an ϵ-neighborhood of the set A and $\mu(A)$ the Lebesgue measure of A. Suppose there is a (discontinuity) set D so that

$$\mu(N_\epsilon(D \cap S_m)) \to 0$$

as $\epsilon \to 0$ for each $m < \infty$. For each $\epsilon' > 0$, let there be an $\epsilon > 0$ so that $|x-y| < \epsilon$ implies $|\hat{f}(x,t) - \hat{f}(y,t)| < \epsilon'$ uniformly in x in bounded regions, provided that $x \notin N_\epsilon(D)$.

Assume (C1). Let Ω denote the sample space. We use the notation $(\Omega, z(t), \mathscr{B}_t, P)$ for the Wiener process on $[0,\infty)$, where $\mathscr{B}_t$ measures $z(s)$, $s \leq t$ and $z(r_2) - z(r_1)$ is independent of $\mathscr{B}_t$ for $t \leq r_1 \leq r_2$, and P is the measure on all the $\mathscr{B}_t$. We say that $z(t)$ is a Wiener process on $(\Omega, \mathscr{B}_t, P)$. Let $x(t)$ be the unique solution to the Itô equation (5)

$$(5) \qquad dx = \begin{array}{l} dx^1 = f^1(x)dt \\ \\ dx^2 = f^2(x)dt + \hat{\sigma}(x)dz \end{array}$$

We say that $x(t)$ is an Itô process with respect to $(\Omega, z(t), \mathscr{B}_t, P_x)$, where P_x denotes the probability given that $x(0) = x$ (and E_x denotes the corresponding expectation). E and P denote expectation and probability for functionals of $z(t)$. Define Ω_T as the sample space for $z(t)$, $t \leq T$. Suppose that

$$(6) \qquad \int_0^T |\hat{\sigma}^{-1}(x(t))\hat{f}(x(t))|^2 dt < \infty \qquad \text{w.p.l.}$$

(which is certainly true if $\hat{f}$ is bounded). Define

$$\zeta_0^T(\hat{f}) \equiv \int_0^T \hat{\sigma}^{-1}(x(t))\hat{f}(x(t))dz(t) - \frac{1}{2}\int_0^T |\hat{\sigma}^{-1}(x(t))\hat{f}(x(t))|^2 dt$$

and suppose that

$$(7) \qquad E_x \exp \zeta_0^T(\hat{f}) = 1.$$

((7) holds for all $T < \infty$ if $\hat{f}$ is bounded.) Then the probability

measure $\widetilde{P}^T$ defined by[+]

$$\widetilde{P}^T_x(A) = \int_A \exp \zeta^T_0(f) \cdot P(d\omega)$$

is a measure on the $\mathscr{B}_t$, $t \leq T$. The process $\widetilde{z}(t)$, $t \leq T$

$$\widetilde{z}(t) = z(t) - \int_0^t \hat{\sigma}^{-1}(x(s))\hat{f}(x(s))ds$$

is a Wiener process on $(\Omega_T, \mathscr{B}_t, \widetilde{P}^T_x)$, and the process

$$(8) \qquad dx = \begin{array}{l} f^1(x)dt \\ f^2(x)dt + \hat{f}(x)dt + \hat{\sigma}(x)[dz - \hat{\sigma}^{-1}(x)\hat{f}(x)dt] \end{array}$$

$$= \begin{array}{l} f^1(x)dt \\ f^2(x)dt + \hat{f}(x)dt + \hat{\sigma}(x)d\widetilde{z} \end{array}$$

is an Itô process with respect to $(\Omega_T, \widetilde{z}(t), \mathscr{B}_t, \widetilde{P}^T_x)$. The construction was first done by Girsanov [4], and exploited by Benes [5], Rishel [2] and then Kushner [3], for several control problems. Note the sample space Ω_T, the σ-algebras $\mathscr{B}_t$ and the random variables $x(t)$ for the Wiener process $\widetilde{z}(t)$, and Itô process $(\Omega_T, \widetilde{z}(t), \mathscr{B}_t, \widetilde{P}^T_x)$ are the same as those for the Wiener process $z(t)$ and Itô process (5), for $t \leq T$. Only the measures have been changed. The process (8) is constructed by a transformation of measures on the 'nicer' process (5).

[+]The measure $\widetilde{P}^T_x$ depends on the initial condition of (5), as does the Wiener process $\widetilde{z}(t)$.

The following facts (drawn from [2-4]) about (8) will be needed. <u>Assume that $\hat{f}$ is bounded and that (Cl-3) hold.</u>

(Ol) ([3], Theorem 5). The multivariate distributions of (8) are continuous with respect to the initial condition $x(0)$, (in the sense that the characteristic functions are continuous in $x(0)$).

(O2) ([3], Theorem 2). The solution to (8) is unique in the sense that any two solutions to (8) have the same multivariate distributions.

$$(O3) \qquad \tilde{E}_x^T \sup_{t>s>0} |x(s) - x|^2 \leq K_1 t(1+|x|^2), \quad t \leq T$$

$$\tilde{E}_x^T \sup_{t>s>0} |x(s) - x|^4 \leq K_1 t^2(1+|x|^4), \quad t \leq T$$

where $\tilde{E}_x^T$ is the expectation given $x(0) = x$, and K_1 depends on the bound on $\hat{f}$. The proof of (O3) is close to that of (27) - (28) of [3] Theorem 6. K_1 depends on the bound on $\hat{f}$.

(O4) If the process (5) has a density $p(x; t,y)$, then so does (8) and the density of (8) is any version of ([2], Lemma 1), (boundedness of $\hat{f}$ is not required if (6) - (7) hold) for $t < T$

$$q(x; t,y) = \tilde{E}_x^T[\exp \zeta_0^t(\hat{f})|x(t) = y]p(x; t,y).$$

Also ($\hat{f}$ is not required to be bounded in (O5)).

(O5) ([4], Corollary to Lemma 3). Let $g(\omega)$ be $\mathscr{B}_t$ measurable with $\tilde{E}_x^T|g(\omega)| < \infty$, and $t \leq T$. Then, for $s \leq t \leq T$, w.p.1.

$$\tilde{E}_x^T[g(\omega)\mid \mathscr{B}_s] = E[g(\omega)\exp \zeta_s^t(\hat{f})\mid \mathscr{B}_s].$$

(The equation also holds if $\mathscr{B}_s$ is replaced by any sub σ-algebra of $\mathscr{B}_s$.)

Fix T, and define $\tilde{z}(t)$ and $\tilde{P}_x^T$ by the Girsanov transformation. Write $\tilde{z}(t)$ as $\tilde{z}^T(t)$. Suppose that (6) - (7) hold for a time $T_1 > T$, and define the corresponding Ω_{T_1}, $\tilde{z}^{T_1}(t)$, $\tilde{P}_x^{T_1}$. Then $\tilde{z}^{T_1}(t) = \tilde{z}^T(t)$ for $t \leq T$, and on sets B of $\mathscr{B}_T$ we have $\tilde{P}_x^T(B) = \tilde{P}_x^{T_1}(B)$. This follows from (05) since (χ_B is the characteristic function of the set B)

$$\tilde{P}_x^{T_1}(B) = E_x[E_x(\chi_B \exp \zeta_0^{T_1}(\hat{f})\mid \mathscr{B}_T)]$$

$$= E_x \chi_B \exp \zeta_0^T(\hat{f})[E_x(\exp \zeta_T^{T_1}(\hat{f})\mid \mathscr{B}_T)]$$

$$= E_x \chi_B \exp \zeta_0^T(\hat{f}) = \tilde{P}_x^T(B).$$

Thus $\tilde{P}_x^{T_1}$ is an extension of $\tilde{P}_x^T$. If (6) - (7) hold for each $T < \infty$, we can replace Ω_T by Ω and define a unique measure $\tilde{P}_x$ on all the $\mathscr{B}_t$, $t < \infty$, which will be consistent with the $\tilde{P}_x^T$ on $\mathscr{B}_T$. Then $\tilde{z}(t)$ will be an Itô process with respect to $(\Omega, \mathscr{B}_t, \tilde{P}_x)$, and $(\Omega, \tilde{z}(t), \mathscr{B}_t, \tilde{P}_x)$ an Itô process (for all $t < \infty$). Both (6) - (7) hold for all $T < \infty$ if $\hat{f}$ is bounded. Let $\mathscr{B} = \bigcup_{t > 0} \mathscr{B}_t$.

3. Existence of a Solution to (8) for Unbounded $\hat{f}$

Let $V(x)$ denote a non-negative twice continuously differentiable function which tends to infinity as $|x| \to \infty$. Define $Q_N = \{x: V(x) < N\}$ and let $\hat{f}_N(x) = \hat{f}(x)$ for $x \in Q_N$ and $\hat{f}_N(x) = 0$, $x \notin Q_N$. Define $C_N^T = \{\omega: x(t) \in Q_N, t \in [0,T]\}$. Let $\tilde{\mathcal{L}}$ denote the differential generator of the process (8) and write $\tilde{\mathcal{L}}^N$ for the differential generator when $\hat{f}$ is replaced by $\hat{f}^N$ in (8). Theorem 1 uses a stability idea to prove existence for (8), for all $t < \infty$.

Theorem 1. Assume (C1) and the above conditions on $V(x)$. Let $\tilde{\mathcal{L}}V(x) \leq 0$ for x not in some Q_a, $a < \infty$. Then

$$(9) \qquad E_x \exp \zeta_0^T(\hat{f}) = 1$$

for all $T < \infty$, and

$$\tilde{z}(t) = z(t) - \int_0^t \hat{\sigma}^{-1}(x(s))\hat{f}(x(s))ds$$

is a Wiener process, for all $t < \infty$ with respect to $(\Omega, \mathcal{B}_t, \tilde{P}_x)$. The solution to (8) exists for all $t < \infty$. It is an Itô process with respect to $(\Omega, \tilde{z}(t), \mathcal{B}_t, \tilde{P}_x)$ and, under the additional assumptions (C2-3), it is unique (in the sense that the multivariate distributions of any two solutions are equal) for all $t < \infty$.

Remark. Let $f(y), \sigma(y)$ satisfy (3), (2) locally, and let $\mathcal{L}_1$ denote the differential generator, with coefficients determined by $f(y), \sigma(y)$. If $V(x)$ and $\mathcal{L}_1 V(x)$ have the properties required in Theorem 1, then the proof can be altered to yield existence and uniqueness for the process $dy = f(y)dt + \sigma(y)dz$.

Proof. Let $\hat{f}^N$ replace $\hat{f}$, in (8), where $N > a$. Let $\tilde{P}_x^{N,T}$ denote the transformed measure with $\tilde{P}_x^{N,T}(A) = \int_A \exp \zeta_o^T(\hat{f}^N)dP$ and $\tilde{P}_x^N$ the extension of $\tilde{P}_x^{N,T}$ to the σ-algebra $\mathcal{B}$ on Ω. Write the Wiener process corresponding to $\tilde{P}_x^N$ as $\tilde{z}^N(t)$ (instead of $\tilde{z}(t)$). Then (8) is an Itô process with respect to $(\Omega, \tilde{z}^N(t), \mathcal{B}_t, \tilde{P}_x^N)$. By virtue of (03) (for $x = x(0)$)

(10) $$\tilde{P}_x^N\{ \sup_{t>s>0} |x(s) - x| \geq \epsilon > 0\} \to 0$$

as $t \to 0$, uniformly for x in compact intervals. Also $\tilde{\mathcal{L}}^N V(x) \leq 0$ in$^+$ $Q_N - Q_a - \partial Q_a \equiv Q_{N,a}$. Let τ denote the first exist^{++} time of the path $x(t)$ from $Q_N - Q_a - \partial Q_a$, and $t \cap \tau \equiv \min (t, \tau)$. Then, by Itô's Lemma $\tilde{E}_x^N V(x(t \cap \tau)) - V(x) \leq 0$ for $x \in Q_N - Q_a$. Since

---

$^+ \partial Q_N$ is the boundary of the set Q_N.

$^{++}$If τ is undefined for some path, set $\tau = +\infty$. Note that the exit time $\tau(\omega)$ (as a path function) for $x(t)$ and $\tilde{x}^N(t)$ are the same, but their distributions may differ.

$$\widetilde{E}_x^N V(x(t \cap \tau)) - a \le (N-a)\widetilde{P}_x^N \{x(s) \quad \text{hits} \quad \partial Q_N \quad \text{before} \quad \partial Q_a \quad \text{and}$$

$$\text{leaves} \quad Q_{N,a} \quad \text{in} \quad [0,t]\},$$

we can conclude that

(11) $\quad \widetilde{P}_x^N \{x^N(t) \text{ hits } \partial Q_N \text{ before } \partial Q_a \text{ and leaves } Q_{N,a} \text{ in } [0,T]\} \le \dfrac{V(x)-a}{N-a} \equiv \epsilon_3.$

We will show that for each $\epsilon > 0$, there is an $N < \infty$ so that

(12) $$\widetilde{P}_x^N \{C_N^T\} \ge 1 - \epsilon.$$

Fix $a_1 > a$. Let $x \in \partial Q_a$. There is a $\delta_o > 0$ so that

$$\min_{x \in \partial Q_a, y \in \partial Q_{a_1}} |x-y| \ge \delta_o.$$

Let $A \in C_N^T$. Then, since $\widehat{f}^N(x(t)) = \widehat{f}^M(x(t))$ on $[0,T]$ for $M \ge N$ and $\omega \in C_N^T$, we have

(13) $$\widetilde{P}_x^M(A) = E_x \exp \zeta_o^T(\widehat{f}^M)\chi_A = E_x \exp \zeta_o^T(\widehat{f}^N)\chi_A = \widetilde{P}_x^N(A).$$

(03) implies that

$$\sup_{x \in \partial Q_a} \tilde{P}_x^N \{ \max_{\delta_1 \geq t \geq 0} |x(t) - x| \geq \delta_o \} \leq K_2 \frac{\delta_1^2}{\delta_o^4} = \epsilon_2.$$

But (13) implies that the constant K_2 depends only on the number a_1 and does not depend on N, for $N > a_1$. Thus, we can assume that K_2 does not depend on N.

Let G_N^T denote the event that $x(t)$ goes to ∂Q_a before ∂Q_N (or never leaves $Q_{N,a}$), then takes more time than $T/n \equiv \delta_1$ to reach ∂Q_{a_1}, then returns to ∂Q_a no fewer than $n - 1$ additional times and after each return takes no less than δ_1 to reach ∂Q_{a_1}, before leaving Q_N for the first time. Then $\tilde{P}_x^N \{C_N^T\} \geq \tilde{P}_x^N \{G_N^T\}$ and $\tilde{P}_x^N \{G_N^T\} \geq 1 - n(\epsilon_1 + \epsilon_2) - \epsilon_3$, where

$$\epsilon_1 = \max_{x \in \partial Q_{a_1}} \tilde{P}_x^N \{x(t) \text{ reaches } \partial Q_N \text{ before } \partial Q_a \} \leq \frac{(a_1 - a)}{(N - a)}.$$

Thus, using $\delta_1 = T/n$,

$$\tilde{P}_x^N \{G_N^T\} \geq 1 - n \left(\frac{a_1 - a}{N - a} + \frac{K_2 T^2}{n^2 \delta_o^4} \right) - \frac{V(x) - a}{N - a}$$

and N and n can be chosen so that $\tilde{P}_x^N \{G_N^T\} \geq 1 - \epsilon$.

There is a unique measure $\tilde{P}_x^T$ on $\mathscr{B}_T$ which is consistent with the $\tilde{P}_x^N$ on the sets C_N^T. Furthermore, (the left hand inequality is [4], Lemma 2)

$$1 \geq \widetilde{P}_x^T(\Omega_T) = E_x \exp \zeta_0^T(\hat{f}) \geq E_x \exp \zeta_0^T(\hat{f}^N) \chi_{C_N^T} \geq 1 - \epsilon.$$

Since ϵ is arbitrary, (9) holds, $\widetilde{z}(t)$, $t \leq T$, is a Brownian motion with respect to $(\Omega_T, \mathscr{B}_t, \widetilde{P}_x^T)$ and $x(t)$, $t \leq T$, an Itô process with respect to $(\Omega_T, \widetilde{z}(t), \mathscr{B}_t, \widetilde{P}_x^T)$. Furtheremore, since T is arbitrary, we can replace $t \leq T$ by $t < \infty$ and $\widetilde{P}_x^T$ and Ω_T by $\widetilde{P}_x$ and Ω.

The process (8) is unique in the following sense. Suppose that both $x^i(t)$, $i = 1,2$ satisfy (8). Let $x^{i,N}(t)$ denote the processes which result when $\hat{f}^N$ replaces $\hat{f}$. Suppose that if $x^{i,N}(t) \in Q_N$ for all $t \in [0,T]$, then $x^i(t)$ coincides with $x^{i,N}(t)$ on $[0,T]$. Then the uniqueness of the $x^{i,N}(t)$ (in the sense of multivariate distributions) and the fact that $\widetilde{P}_x^N\{C_N^T\} = \widetilde{P}_x^M\{C_N^T\} \geq 1 - \epsilon$ for $M > N$ (the $\widetilde{P}_x^N$ do not depend on i) imply uniqueness of the $x^i(t)$ in the sense of multivariate distributions. Q.E.D.

Remark. Lemma 7 of [4] would appear to yield existence for a large class of unbounded $\hat{f}$. But an examination of the proof shows that its content is the following. Let processes (5) and (8) exist with respect to some Wiener process, with (5) being unique, and $\int_0^T |\hat{\sigma}^{-1}(x(t))\hat{f}(x(t))|^2 dt < \infty$ w.p.1, where $x(t)$ is the solution to (5). Under some minor subsidiary condition, it is proved that

$$E_x \exp \zeta_0^T(\hat{f}) = 1$$

where the expectation corresponds to (5). Then (8) can be obtained

by a Girsanov transformation from (5). But both the square

integrability property and existence for (8) must be established

first. But these properties are essentially the desired result.

3. Markov Properties of (8)

Write (C4): In each compact x set, there is an $\alpha > 1$

and $M < \infty$ so that

$$\int p^{\alpha}(x; t,y) \le M < \infty.$$

Theorem 2. Assume (C1) - (C3) and the condition on V and

$\tilde{\mathcal{L}}V$ of Theorem 1. Then the process (8) is a strong Markov process.

If (C4) holds, for some $\alpha > 1$, (8) is a strong Feller process.

Proof. T he terminology of Theorem 1 will be used. By

Theorem 1, the process is defined on the time interval $[0,\infty)$, and

has continuous paths w.p.1.

First, we prove that (8) is a Markov process. Let $\mathcal{B}_t^x \subset \mathcal{B}_t$

measure $x(s)$, $s \le t$. Define the transition function $\tilde{P}_x(x; t,A) =$

$\tilde{P}_x\{x(t) \in A\}$. Since the right hand term of

$$\tilde{P}_x\{x(t) \in A\} = E_x \chi_{\{x(t) \in A\}} \exp \zeta_0^t(\hat{f})$$

is a Borel measurable function of x, so is $\tilde{P}(x; t,A)$ for each $A \in \mathscr{B}_t^x$. <u>Now assume that</u> $\hat{f}^N$ <u>replaces</u> $\hat{f}$. The Chapman-Kolmogorov equation holds since, by (05) and the fact that (5) is a Markov process,

$$\tilde{E}_x^N[\chi_{\{x(t+s)\in A\}} \mid \mathscr{B}_s^x] = E_x[\chi_{\{x(t+s)\in A\}} \exp \zeta_s^{s+t}(\hat{f}^N) \mid \mathscr{B}_s^x]$$

$$= E_{x(s)}[\chi_{\{x(t)\in A\}} \exp \zeta_0^t(\hat{f}^N)] = \tilde{P}^N(x(s);t,A)$$

w.p.1. Thus by the definition Dynkin [6, Chapter 3], $x^N(t)$ (the Itô process on $(\Omega, \tilde{z}^N(t), \mathscr{B}_t, \tilde{P}_x^N)$ corresponding to the use of $\hat{f}^N$) is a Markov process.

The σ-algebras $\mathscr{B}_t^x$ also measure (8). The measure $\tilde{P}_x$ for the unbounded $\hat{f}$, has the correct conditioning properties since, by (05) and the dominated convergence theorem,

$$\tilde{E}_x[\chi_{\{x(t+s)\in A\}} \chi_{\{C_{t+s}^N\}} \mid \mathscr{B}_s^x]$$

$$= E_x[\chi_{\{x(t+s)\in A\}} \chi_{\{C_{t+s}^N\}} \exp \zeta_s^{t+s}(\hat{f}) \mid \mathscr{B}_s^x]$$

$$\to E_x[\chi_{\{x(t+s)\in A\}} \exp \zeta_s^{t+s}(\hat{f}) \mid \mathscr{B}_s^x] =$$

$$= E_{x(s)}[\chi_{\{x(t)\in A\}} \exp \zeta_0^t(\hat{f})] = \tilde{P}(x(s);t,A)$$

w.p.1. Then, by the definition [6, Chapter 3], (8) is a Markov process.

(8) is a Feller[+] process, hence a strong Markov process [6, Theorem 3.10]. The proof is omitted. The proof of the stronger 'strong' Feller[++] property will be given next, under the additional condition (C4). Let (C4) hold.

Supposing that (8) is a strong Feller process if $\hat{f}$ is bounded, we show that it is a strong Feller process for unbounded $\hat{f}$. Let $g(\cdot)$ be bounded and measurable. Then $\tilde{E}_x^N g(x(t)) \equiv G^N(x)$ is continuous in x, for $t > 0$. Write $G(x) = \tilde{E}_x g(x(t))$. Then

$$|G(x) - G^N(x)| \leq \max_x |g(x)| \cdot [\tilde{P}_x\{\Omega - C_N^T\} + \tilde{P}_x^N\{\Omega - C_N^T\}] \to 0 \quad \text{as} \quad N \to \infty,$$

uniformly in any compact x set. Thus, $G(x)$, being the uniform limit of continuous functions, is continuous.

Finally, suppose $\hat{f}$ is bounded and (C4) holds. Reproducing an argument of Rishel [2], we show for each compact x set, there is a $\beta > 1$ and $M < \infty$ so that (q is the density of (8) - see (O4))

(14) $$\int q^\beta(x; t,y) \, dy \leq M_1 < \infty.$$

[+]A process $x(t)$ is a Feller process if $E_x f(x(t))$ is a continuous function of x, if $f(x)$ is continuous and bounded.

[++]$x(t)$ is a strong Feller process if $E_x f(x(t))$ is continuous in x for any bounded Borel function $f(x)$ and $t > 0$.

Define $r(x; t,y) \equiv \tilde{E}_x[\exp \zeta_0^t(\hat{f}) | x(t) = y]$. Let $m^{-1} + n^{-1} = 1$, and note that, for any $\rho > 1$ and compact x set, there is an $N_\rho < \infty$ so that $\tilde{E}_x \exp \rho \zeta_0^t(\hat{f}) \leq N_\rho$ ([4], Lemma 1). Let $\beta > \beta_1$, $\beta > 1$. By Holder's inequality

$$\int p^\beta(x; t,y) r^\beta(x; t,y) = \int p^{\beta_1}(x; t,y) r^\beta(x; t,y) p^{\beta-\beta_1}(x; t,y) dy$$

$$\leq [\int p^{\beta_1 n}(x; t,y) r^{\beta n}(x; t,y) dy]^{1/n} [\int p^{(\beta-\beta_1)m}(x; t,y) dy]^{1/m}.$$

We can choose $\beta > 1$, $\beta > \beta_1$, m, n and $\rho > 1$ so that $(\beta-\beta_1)m = \alpha$, $\beta n = \rho$, $\beta_1 n = 1$, which , together with (C4), proves (14). Equation (14) implies that, as x varies in any compact set, the family $q(x; t,y)$ of functions of y is uniformly integrable. This, together with the con-tinuity (in x) of $\tilde{P}(x; t,(-\infty,a))$ for any vector a (recall that there is a density) implies that $\tilde{P}(x; t,A)$ is continuous in x for any Borel set A, which implies, in turn, the strong Feller property. For more detail, note that the boundary of any rectangle in the range space of $x(t)$ has zero probability, and that $\tilde{P}(x; t,A)$ is continuous in x if A is the sum of rectangles (open or closed). Let $\tilde{P}(x; t,A_j)$ be continuous in x for a collection of sets A_j, which increase monotonically to A

$$\tilde{P}(x; t,A) = \int_{A_j} q(x; t,y) dy + \int_{A-A_j} q(x; t,y) dy.$$

The second integral goes to zero as $j \to \infty$ uniformly in x in any

compact set, by the uniform integrability of $q(x; t,y)$. Since the first integral is continuous, so is the uniform limit $\tilde{P}(x; t,A)$. Q.E.D.

4. The Invariant Measure, and the Asymptotic Properties of the Measures of (8)

In [8], under the conditions (D1) - (D5), Khasminskii proved the existence of a unique σ-finite invariant measure for a process $x(t)$ with a stationary transition function $\tilde{P}(x; t,A)$ under the conditions (D1-5).

(D1) For any ϵ neighborhood $N_\epsilon(x)$ of x, $1 - P(x; t,N_\epsilon(x)) = o(t)$ uniformly in x in any compact set.

(D2) The process is a strong Markov and strong Feller process.

(D3) $\tilde{P}(x; t,U) > 0$ for all open sets U and $t > 0$.

(D4) The paths are continuous w.p.1.

(D5) The process is recurrent. (There is some compact set K and a random time $\tau < \infty$ w.p.1. so that $x(\tau) \in K$ w.p.1., for each initial condition.)

In [9], Kushner applied the result in [8] to obtain a sufficient condition for the convergence of the measures of class of diffusions to a unique invariant measure. Theorem 3 includes the prior result as a special case. Zakai [10] has treated the invariant measure problem for a class of diffusions satisfying (2) - (3), using

a general method of Benes [11]. A similar problem is treated in Elliot [12]. Elliot's method involves a condition on a Lie algebra generated by certain functions of the diffusion coefficients, which is hard to check in special cases. The result of Benes [11] (concerning only existence of an invariant measure) uses the condition that

$$\lim_{|x| \to \infty} P(x; t,K) \to 0 \quad \text{for all compact sets} \quad K. \quad \text{This would not always}$$

hold under our conditions. E.g., the solution to $\dot{x} + x^3 = 0$, reaches $x = 1$ in a time that is bounded as $x(0) \to \infty$, and we would expect a pect a similar result for $dx = -x^3 dt + \sigma dz$.

Theorem 3. Assume (C1) - (C4), and the conditions on $V(\cdot)$ in Theorem 1, except let $\widetilde{\mathscr{L}}V(x) \leq -\epsilon < 0$ outside of Q_a. Let (5) have a nowhere-zero density, for each initial condition x. Then (8) has a unique invariant measure $Q(\cdot)$ and $\widetilde{P}(x; t,A) \to Q(A)$ as $t \to \infty$ for any x. Both $\widetilde{P}(x; t,A)$ and $Q(A)$ have nowhere zero densities.

Remark. Theorem 3 only deals with invariant measures, but almost all of stability results in [1] can be carried over to the problem with discontinuous drift terms.

Proof. The second inequality of (O3) implies (D1) for bounded $\hat{f}$, and, hence, for the processes $x^N(t)$. But, if (D1) holds for each $x^N(t)$, it holds for (8). (D2) is proved in Theorem 2. Since $\widetilde{E}_x[\exp \zeta_0^t(\hat{f})|x(t) = y] > 0$ w.p.1. and $p(x; t,y) > 0$ for

y by assumption, $q(x;\ t,y)$ (the density for $\widetilde{P}(x;\ t,A)$) is positive for almost all y (Lebesgue measure). This implies (D3). (D4) is a consequence of Theorem 1. (D5) is a consequence of $\widetilde{\mathscr{L}}V(x) \leq -\epsilon < 0$ for all large x. (See Theorem 4 in [9]). Indeed, the average time to leave the set $Q_N - Q_a - \partial Q_a$ (for $x(0) = x$) is bounded above by $(V(x) - a)/\epsilon < \infty$. This together with (11) gives (D5). Thus all (D1-5) hold.

$Q(A)$ satisfies

$$Q(A) = \int Q(dx)\widetilde{P}(x;\ t,A)$$
$$= \int_A du \int Q(dx)q(x;\ t,u).$$

Thus $Q(A) > 0$ for all sets A of positive Lebesgue measure and has density $\int Q(dx)q(x;\ t,u)$, which must be positive almost everywhere.

For a process with a transition density and a unique invariant measure $Q(\cdot)$ with a nowhere zero density, Doob [7, Theorem 5] proves that $\widetilde{P}(x;\ t,A) \to Q(A)$ as $t \to \infty$ for any x. Q.E.D.

References

[1]   H. J. Kushner, Stochastic Stability and Control, Academic
        Press, New York, 1967.

[2]   R. W. Rishel, "Weak solutions of the partial differential
        equations of dynamic programming", submitted to the SIAM
        Journal on Control.

[3]   H. J. Kushner, "Approximations to and local properties of
        diffusions with discontinuous controls", submitted to the
        SIAM Journal on Control.

[4]   I. V. Girsanov, "On transforming a certain class of stochastic
        processes by absolutely continuous substitution of measures",
        Theory of Probability and Its Applications, 5, 1960, 285-301,
        translation.

[5]   V. E. Benes, "The existence of optimal strategies based on
        specified information for a class of stochastic decision
        problems", SIAM Journal on Control, 1970.

[6]   E. B. Dynkin, Markov Processes, Springer, Berlin, 1965,
        Translation from 1962 Russian Edition.

[7]   J. L. Doob, "Asymptotic properties of Markov transition
        probabilities", Transactions of the American Mathematical
        Society 3, 1948, 393-421.

[8]   R. Z. Khasminskii, "Ergodic properties of recurrent diffusion
        processes and stabilization of the solution to the Cauchy
        problem for parabolic equations", Theory of Probability and Its
        Applications 9, (1960), 179-196.

[9]   H. J. Kushner, "The cauchy problem for a class of degenerate
       parabolic equations, and asymptotic properties of related
       diffusion processes", J. Differential Equations $\underline{6}$, 1969,
       209-231.

[10]   M. Zakai, "A Lyapunov criteria for the existence of stationary
       probability distributions for systems perturbed by noise",
       SIAM Journal on Control 7, 1969, 390-397.

[11]   V. E. Benes, "Finite regular invariant measures for Feller
       processes", Journal of Applied Probability $\underline{5}$(April) 1968.

[12]   D. L. Elliot, "Behavior of controllable nonlinear systems
       driven by white noise", Second Hawaiian Conference on System
       Science, January 1969. Dissertation, UCLA Engrg. Dept., 1969.